DOCUMENT RESUME

ED 371 813 JC 940 422

AUTHOR	Pfeiffer, Jay J.
TITLE	Student Follow-Up Using Automated Record Linkage Techniques: Lessons from Florida's Education and Training Placement Information Program (FETPIP).
PUB DATE	May 94
NOTE	22p.
PUB TYPE	Guides - Non-Classroom Use (055) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Agency Cooperation; Cooperative Planning; *Databases; *Employment Patterns; *Information Retrieval; *Outcomes of Education; Postsecondary Education; Program Development; Program Implementation; *Vocational Followup
IDENTIFIERS	*Education and Training Placement Info Prog FL; Florida

ABSTRACT

Florida's Education and Training Placement Information Program (FETPIP) is a statewide system linking the administrative databases of certain state and federal agencies to collect follow-up data on former students or program participants. The databases that are collected include those of the Florida Department of Corrections; Florida Department of Education, including the Board of Regents, State Board of Community College, Division of Public Schools, and Office of Student Financial Assistance; U.S. Department of Defense; U.S. Office of Personnel Management; U.S. Postal Service; Florida Department of Administration; and Florida Department of Labor and Employment Security. FETPIP provides follow-up services to 75 programs and organizations on more than 1.8 million former students, participants, ex-convicts and trainees. This article provides guidance to states who may be considering similar programs in the areas of: (1) collecting data by computer matching; (2) recognizing the ownership of individual data; (3) taking the initial steps (i.e., determine interest and receptivity, determine conditions under which administrative records can be linked, review wage record and student/participant record structure and content, involve the agency responsible for administering the state's unemployment insurance program, and determine who does what); (4) developing interagency agreements; (5) carrying out the interagency agreements; and (6) cultivating ideas for the future direction of the system. (KP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 371 813

Student Follow Up Using Automated Record Linkage Techniques: Lessons from Florida's Education and Training Placement Information Program.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J. Pfeiffer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Jay J. Pfeiffer

JC 940 422

STUDENT FOLLOW-UP USING AUTOMATED RECORD LINKAGE TECHNIQUES: LESSONS FROM FLORIDA'S EDUCATION AND TRAINING PLACEMENT INFORMATION PROGRAM (FETPIP)

Submitted by:
Jay J. Pfeiffer, Director
Florida Education and Training Placement Information Program
Florida Department of Education

Abstract

Florida's Education and Training Placement Information Program (FETPIP) is a statewide system that relies on the administrative data bases of certain state and federal agencies to collect follow-up data on former students or program participants. The program links the records of the former students/program participants to the administrative data bases to obtain follow-up data. Data that are collected include civilian employment, continuing postsecondary education, and military enlistment. Extensive automation is used in collecting and reporting data. Results are provided to state and local counselors, teachers, and education/training policy makers for planning, evaluation, and career guidance purposes. Data are a part of education and training accountability programs at all levels in the state.

This article will describe the program's operation and provide guidelines for those considering such an approach.

Introduction

Those who administer education, employment, and training programs have, for the past fifteen years, been continuously asked to accomplish more with less funding. They have been asked to focus more and more on serving clients whose education, training, and support service needs are extensive. Furthermore, they have been required to document the consequences of their activities as a condition of continued funding. However, the traditional methods used for such documentation have been expensive and their results inconclusive.

During the early 1980s, Florida's legislative and executive branches conducted extensive reviews of the state's education, employment, and training programs with a particular focus on vocational education. The reviews included all aspects of program activities from planning, operating, and coordinating to evaluating. Among the findings were several that indicated that the information available to support program activities were inadequate, especially those dealing with labor market analysis and program evaluation.

There were a myriad of recommendations that were posed as a result of the various studies. Among them was one that suggested using employer payroll data.

collected through the Florida unemployment compensation program as a tool in evaluating programs and services. As a result of this suggestion, the legislature created a research effort called the "Occupational Identifier Project" in 1984. At the same time, the legislature enacted a performance requirement for secondary and postsecondary vocational education that stipulated that 70% of any program's completers had to be placed. "Placement" was defined as employment in a job related to training, continuing postsecondary education, or enlistment in uniformed military service of the United States. When enacted, data for this performance requirement could be collected through local mail or telephone surveys or through use of data from the Occupational Identifier Project. In 1988 as a result of the Project's accomplishments, the legislature enacted Section 229.8075 of Florida Statutes creating the Florida Education and Training Placement Information Program. FETPIP then became the primary resource for documenting compliance with the vocational education performance requirement.

Since 1988, FETPIP has both expanded its follow-up data coverage and its services to education and training organizations.

Current FETPIP Operations

FETPIP provides follow-up services to a variety of "Applications". The term "applications" refers to the programs and organizations for whom follow-up data collection services are provided. Currently there are 75 "applications". They include all Job Training Partnership Act programs, Project Independence participants (Florida's version of the Federal JOBS program for welfare recipients), all public school system high school graduates and dropouts, all community college associate degree students, all secondary and postsecondary vocational students, all state university system graduates, adult education and GED students, selected private vocational school students, all unemployment insurance benefit recipients, and all correctional system releases. Smaller operations such as adult migrant education, blind services, apprenticeship, and certain scholarship recipient groups are served as well. In cooperation with several participating agencies, FETPIP has designated selected groups of former students for longitudinal data collection.

Organizations representing each "application" provide FETPIP with individual student or participant files from their management information system units. The files include individual identifiers (name, social security number) as well as demographic, socio-economic, and programmatic data.

This year the applications account for more than 1.8 million former students, clients, participants, ex-convicts, and trainees. Follow-up data for applications representing these people are currently being collected.

The initial step in FETPIP's processing is to screen SSNs to assure that each reported SSN meets assignment parameters used by the U.S. Social Security Administration and that each number has in fact been assigned (FETPIP receives quarterly updates from the Regional Office of the U.S. Social Security Administration).

Invalid SSNs are eliminated from FETPIP's processing routines.

For the current fiscal year, records containing validated SSNs are being electronically linked with the State Department of Corrections (for new incarcerations and recidivism), the State Department of Education (for former students/participants enrolled in public postsecondary education at any level and private college/university enrollments in Florida). Additional linkages are being conducted through agreements with the U.S. Department of Defense (military enlistments), the Federal Office of Personnel Management (federal career service employment), U.S. Postal Service (postal career service employment), the State Department of Administration (Florida career service employment), and the Florida Department of Labor and Employment Security (employment through Unemployment Insurance wage records).

The following list summarizes the data items that are obtained each time there is a match between the various data bases.

FLORIDA DEPARTMENT OF CORRECTIONS: This linkage primarily focuses on releasees from the state prison population to determine if they had returned to prison during a defined period after release. The current period runs from July 1990 to January 1994 for those released during 1990-91, July 1991 to January 1994 for those released during 1991-92, and July 1992 through January 1994 for those released during 1992-93.

The data include offense, work assignment, location, services, and sentencing data items.

FLORIDA DEPARTMENT OF EDUCATION: Records are linked to four separate management information system units within the Department for a determination of persons continuing their educations in a postsecondary environment. For this year, the records are linked to 1993 fall enrollment records.

The Board of Regents: Linked records included institutional identifiers and declared major codes for each of Florida's nine public four-year universities.

The State Board of Community Colleges: Records included institutional identifiers and major codes for the 28 public community colleges.

The Division of Public Schools: Linked records included data for postsecondary adult and vocational education. Institutions, school districts, and program area data items are collected.

The Office of Student Financial Assistance: Linked records will identify Florida residents pursuing postsecondary education in any of Florida's private colleges or universities (residents receive a voucher when they do so, the voucher program is administered through this office).

THE U.S. DEPARTMENT OF DEFENSE: Linked records will identify former students/trainees who enlisted in military service between July 1993 and March 1994.

Data elements include the person's rank/paygrade, Military Occupational Specialty, and Branch of Service.

THE U.S. OFFICE OF PERSONNEL MANAGEMENT : Linked records will indicate former students/trainees in the employ of the federal government between October and December 1993. Data elements include branch of government and location, federal job classification, and pay grade.

THE U.S. POSTAL SERVICE: For former students who obtained jobs in the Postal Service between October and December 1993, data will be collected that indicate the Postal Job Classification, Pay grade, and location of employment.,

THE FLORIDA DEPARTMENT OF ADMINISTRATION: For former students employed in a state government agency, public university, or the state/local judiciary system between October and December 1993, data will be collected that indicates job titles, agency, earnings, and job location.

THE FLORIDA DEPARTMENT OF LABOR AND EMPLOYMENT SECURITY: The files that are accessed here are a part of the wage report system that is used to manage the state unemployment compensation program. These reports are basically quarterly employer payrolls from throughout Florida. This year's efforts focus on the October - December 1993 quarter and the four quarters immediately preceding it. Where there is a "hit", data are collected that indicates the name and address of the employer, the Standard Industrial Classification of the employer, the total number of employees in that establishment for the reporting period, the number of weeks worked by the individuals found, and their reported earnings for the period.

Armed with the Department of Labor's employment data, FETPIP will contact about 25,000 employers to determine the occupations and county locations of students that were found in their employ. FETPIP refers to this as the employer "sample", though this is not a sample in the statistical sense of the word. Basically, employers in this "sample" are selected because they employed students who completed or graduated from education or training programs where occupational information might be a critical element of program evaluation. For the past seven years FETPIP has been able to maintain an 86% response rate from employers.

This year's employer inquiry focuses on employees hired during the October-December 1993 period. Occupational data coded to the Occupational Employment Statistics Survey-based (OES) code will be obtained. In responding, employers will have the option of: 1. using an industry-specific occupational dictionary with OES codes that will be provided to them; or, 2. using their own personnel system job codes. When FETPIP receives responses using personnel codes, the program's data entry apparatus includes an automated occupational thesaurus that matches employer job titles to the OES Codes.

Once data are collected and reviewed, a variety of reports and files are provided

to the organizations represented by the applications. The general mode of operation is to provide data with technical assistance and advise to such applications.

The data that are collected serve many purposes. They are used in conjunction with State Occupational Information Systems as a compliment to occupational supply and demand data. They are used to meet increasing demands for consumer information. They are also used as a part of program accountability and evaluation processes.

Traditional methods of conducting follow-up studies include mail and/or telephone interviews with former students or participants. Such methods are generally expensive and yield low response rates.[1] It is estimated that FETPIP has saved community colleges and school districts $3.1 million in annual expenses related to conducting traditional follow-up in each year since its inception.[2]

FETPIP's annual budget is approximately $320,000. If the program's matching experience continues at past rates, it will successfully match about half of the people on whom data are being sought.[3] This would mean about 900,000 "completed" follow-ups at a cost of about thirty-six cents for each one. Estimates for completed traditional follow-ups in the JTPA system for similar data are $19.00 each (NCEP, 1992, p.2).[4]

In the 1988-89 fiscal year, FETPIP reached its current budget level. At that time it was following up on about 200,000 former vocational education students with a match rate of 75% (FETPIP, 1990). Unit costs at that time were about $2.13 per matched student. This suggests that the approach is not only less costly than traditional methods, but that as the program expands it can effectively reduce unit costs. At the same time as it expands, it introduces savings into other sectors of the education, employment, and training program system by reducing their follow-up data collection costs.

FETPIP represents a data collection resource that is capable of providing comprehensive employment data on a near universe of former students and program participants at relatively low cost. Individual program, demographic, and socioeconomic characteristics may be electronically combined with information that describes employment characteristics. The employment information includes identifying the type of employer, certain geographic dynamics of employment, certain longitudinal aspects of employment, and earnings. The employment data are available on a timely basis. They are consistently defined and collected over time. They are based on standard definitions that are used to drive other information resources. The wage report data bases are comprehensive in that for most states they come close to representing a quarterly census of wage and salary workers and their employers.

There is increasing interest in conducting follow-up through unemployment insurance wage reports and other data bases. There are many states that are doing follow-up for former vocational education students or JTPA participants through the use of the wage records. An additional few states have begun pursuing a comprehensive follow-up approach such as that practiced by FETPIP. They include Texas, Oregon, North Carolina, North Dakota and Indiana. Several other states including these have

joined together to form an interstate consortium to develop data handling and analysis guidelines for follow-up data collections through record linkage techniques. Consortium states have set out objectives that include establishing a clearinghouse, providing other states with technical assistance where necessary, and reacting to legislation or other events that have potential impact on state activities related to follow-up using wage records.[5] Federal law has prompted several national efforts to examine uses of the wage record resource for follow-up and economic research.[6]

With this interest and growing demands for accountability, readers may be tempted to forge ahead and pursue the collection of follow-up data via wage records on former students, trainees, and/or program participants. A data resource is presented that promises to provide timely and accurate employment and earnings information for a substantial number of former students at a low cost. Consistent statewide definitions of terms are assured. Teachers and administrators in a variety of education and training operations are relieved of the burden and expense of collecting follow-up data.

The purpose of this article is to instill those who are interested in pursuing this resource with a sense of propriety and caution. A lack of circumspection in this endeavor could have repercussions that could limit access to the unemployment insurance wage reports and other resources in the future.

Collecting Data by Computer Matching

> *"...the term 'matching program' means any computerized comparison of two or more automated systems of records or a system of records with non-federal records for the purpose of - establishing or verifying the eligibility of, or continuing compliance with statutory and regulatory requirements by, applicants for, recipients of, participants in, or providers of services with respect to, cash or in-kind assistance or payments under federal benefit programs, or recouping payments or delinquent debts under such federal benefits programs..."*
>
> *"...the term, 'matching program'... does not include - matches performed to support any research or statistical project, the specific data of which may not be used to make decisions concerning the rights, benefits, or privileges of specific individuals..."*
>
> Subsection (a) of Section 552a U.S. Code
> (as amended by the Computer Matching and Privacy Protection Act of 1988)

There is an important distinction in the definitions cited above between a "matching program" and a research or statistical project. While the term "matching program" is frequently used when referring to the subject of this article, it is clear that what is referred to does not meet the "legal definition". Because the process involves linking certain data in two or more data bases through a common data element, in this case, Social Security

Numbers, the term "record linkage" is appropriate.

It is important that readers wishing to pursue data collection through "record linkage" carefully observe this distinction in negotiating arrangements with their unemployment insurance or other data resource counterparts. Purposes associated with research or program evaluation should not be mixed with those that target individuals for sanctions or specific actions.

The collection of data by electronically linking administrative data bases as a means of supporting statistical analyses is a relatively new phenomenon. Its use for vocational education or JTPA follow-up is but one of several applications that have been and are being examined using the technique. It has been used in health and vital statistics by the Center for Disease Control, U.S. Census to Internal Revenue Service master files, enhancements from the U.S. Survey of Income and Program Participation and economic data, and a myriad of others.[7] The Washington Statistical Society, the Federal Commission on Statistical Methodology, and the U.S. Bureau of the Census have periodically cosponsored workshops to help define and refine the technique as well as to share experiences among researchers. While it is a legitimate means of data collection, and it can support the aims of this student follow-up, it is not like a survey process.

With a survey process, the designer/administrator controls much of the process. Questions or terms are carefully selected and designed to elicit clearly defined answers. A sampling design is chosen that best represents a selected population given certain response expectations. The survey is conducted in a manner that maximizes a response level. When data are collected by linking several data bases from different agencies, however, much of this type of quality control is lost.

One is dependent on controls that may or may not be applied by others. For example, the accuracy of wage report data requires that employers accurately record employee identification and payroll information. It also requires that the employer data are entered accurately when received by the unemployment insurance agency. The assignment of Standard Industrial Classification Codes to employers must be accurate as well. Similarly, the accuracy of student data to be used in a record linkage program requires that Social Security Numbers be collected, accurately recorded, verified, and properly transmitted. Other student level information such as demography, socio - economic characteristics, program characteristics, etc. must be faithfully represented.

To retain some modicum of quality control, those interested in collecting follow - up data by linking student records to the unemployment insurance wage report or other records must have a clear understanding of the data sets involved. They must know how the data are originally collected and recorded, how they are processed, how they are defined, and what they represent. As the record linkage activity proceeds, any anomalies or problems should be brought to the attention of the affected agency for clarification or resolution. This should be accomplished in a helpful fashion.

Recognizing the Ownership of Individual Data

Among the things that each of us "own" is information about ourselves. Some of this information might be bandied about freely. Some might be protected vigorously because of its personal nature. Because the information is privately owned, it is a matter of personal choice as to whether it is bandied freely or protected. When an organization, whether it be a school, employer, or an unemployment insurance agency, requests access to this personal property, there is an implicit (and often explicit) agreement that what is being provided remains personal - not organizational - property. The Privacy Act of 1974 (as amended) canonizes certain aspects of this agreement.

The Privacy Act requires that federal agencies collect and maintain only that personal information which is relevant to the lawful purposes of the agency. When personal information is requested from individuals, the Act requires that they be told whether the information must be provided as a requirement of law or is being solicited on a voluntary basis. They are to be told what the information will be used for and how it will be protected. The agency must refrain from disclosing the personal information that is collected to agency employees except in the performance of their lawful duties. The agency must not make personally-identifiable information available outside of the agency without the consent of affected individuals unless it is for certain defined purposes.[8]

The laws that govern the collection of personal information by state unemployment insurance agencies or public educational institutions are founded on these types of principals. They collect only the data that are necessary to fulfill their prescribed information needs.

State unemployment insurance laws allow administering agencies to collect specific data that are necessary for managing the state's unemployment insurance program. The collection of student follow-up data is not an exercise directly connected with managing unemployment insurance.[9]

The laws also require that the agencies be very protective of the personal information that is in effect, loaned to them by employers. They will be correctly reluctant to loan it to someone else without taking some precautions.

Taking the Initial Steps

In many states, the agency interested in pursuing the collection of follow-up data is neither the agency that collects unemployment insurance wage reports nor that which collects student data. This agency, the "match maker", will have to reach some form of legally-binding agreement before data can be collected. However, there are preparatory steps to be taken before an agreement can be drafted, negotiated, and consummated.

Step One: Determine interest and receptivity. This may occur in several ways. In Florida, the Governor's staff prompted an initial interest in student follow-up using wage records through its 1982 vocational education study. This interest ultimately took the form

of legislation. In this situation the interest was prompted by higher authority.

In Texas, a local community college had experience in doing follow-up with wage reports through its own initiative. This activity initially expanded to include a consortium of colleges. In a special feasibility study examining potential uses of the wage report, it was recommended that a statewide, comprehensive follow-up system be developed. In this case, an education or training agency already had experience with follow-up by linking student records to wage reports. State interests, through the State Occupational Information Coordinating Committee, were able to build upon the local experience to include additional institutions and agencies.

In North Dakota, a state level meeting was sponsored by the State Occupational Information Coordinating Committee with representatives of agencies that would have potential interest in the use of data collected through the wage report. The meetings were structured to introduce and discuss the concept in detail. Following the initial meeting, a series of intimate meetings were carried out with each agency to build on the initial interest. In this case, interest was prompted by open discussion of the idea of a student record/wage record type of follow-up linkage.

Step Two: Determine conditions under which administrative records can be linked. An additional initial step will be to determine whether or not and under what conditions Social Security Numbers can be used for the type of follow-up discussed in this article. It may be judicious to involve the legal staff of the "match maker" agency early so that they can be of assistance later. If there are state laws or interpretations of federal laws that prohibit such uses, those who are interested in this type of follow-up will have to judge how best to proceed. One approach might be to develop an issue paper to circulate among officials in the executive or legislative branches to determine if there is sufficient interest to pursue appropriate legislative remedies. This interest may be cultivated as officials learn of related activities at the national level and in other states. It may also be cultivated by the notion of developing information that assists in holding education and training programs accountable for results at relatively low cost.

Step Three: Review wage record and student/participant record structure and content. If follow-up by linking records through Social Security Numbers is a permissible activity, another initial step is to determine whether or not individually identifiable student records include Social Security Numbers at any level (i.e., the institutional, regional, or state level) or even at any location. If they are included anywhere, it would be useful to determine if the files are electronically accessible. There are two areas of activity that should accompany this step. One is technical and relates to management information and the manner in which data are defined, stored, processed, and accessed. Work in this area should also be directed to defining data elements. Another area of activity is promotional and is designed to continue encouraging participation at the appropriate level or location.[10] The promotional area of work should also be directed to defining exactly what type of reports are desired from student follow-up with administrators and policy makers at the selected level.

In some cases Social Security Numbers will not be a part of the repertoire in the student data files. An institution's administrators may be convinced to begin collecting the numbers either in future school years or as a supplemental activity in the current school year. If this occurs, the "match maker" should work with administrators to develop a collection process that exceeds the requirements of the Privacy Protection Act with respect to informing students and their parents and to providing adequate controls that secure and protect the information.

Step Four: Involve the agency responsible for administering the state's unemployment insurance program. An initial step will be to determine the receptiveness of the UI agency toward a follow-up effort. Like the student record inquiry, this interaction should have technical as well as promotional aspects.

All states use wage reports in some form as a part of the employer tax/eligibility determination process for unemployment insurance. Therefore, there will be an automated system with certain employer and employee information including the social security number. The technical part of the inquiry will have to be directed to defining information definitions, flows, and content. It should also be directed to determining how the wage report is being used by other agencies.

The wage report is currently being used in most states for several purposes other than administering the unemployment insurance program. For example, because it is comprehensive, it is a major information resource used by state labor market information units (also referred to as Research and Statistics or Research and Analysis units) in state employment security agencies. This is the reason that the Standard Industrial Classification is assigned to each employer in the system. While such use is not strictly related to the unemployment insurance program, it is considered an "in the family" use because of the unit's affiliation with the state employment security agency.

There are a number of "out of family" users as well. These other users include those who match individual Social Security Numbers against the wage report to verify eligibility for certain programs (such as JTPA and AFDC), garner wages for non payment of federal student loans or child support, and identify subjects of interest to law enforcement agencies.

Other agencies or organizations conducting follow-up studies might include those which administer the federal JOBS program or the JTPA program. They could include university economic or social research units. They also might include local school districts, vocational-technical education institutions, or community colleges.

Where there are other units doing follow-up studies, the "match maker" may have several options in obtaining the unemployment insurance wage report data. These options could include a dovetail type of process, where student records are combined with other types of individual records when they are submitted to the unemployment insurance agency by the "other unit". An additional option could be to act as a coordinator by working to combine multiple interactions with the UI wage report data base

into a single interaction. Where these efforts can be consolidated, an actual service is being provided to the unemployment insurance agency as well as to the education and training agencies. Regardless of whether there are other follow-up studies being conducted, it will be desirable to enter into a record linkage agreement with the unemployment insurance agency alone.

Step Five: Determine who does what. There are several facets to the "Who does what?" question. One is to assign responsibility for coordinating the receipt of records from education, employment, and training organizations that will be used to link to the wage report. This organizational entity has been referred to as the "match maker" in this article. This organization could be the State Occupational Information Coordinating Committee as in Texas and North Dakota, the Employment Security Agency as in Oregon and North Carolina, or a special unit within the Department of Education as in Florida.

Responsibility for actually conducting computer linkages will have to be assigned. In Florida, FETPIP coordinates the collection of student/participant records from participating agencies and provides a computer tape containing the records to the unemployment insurance agency where the actual linkage occurs. Linkages with the federal government for employment data and to the university system for enrollment data occur similarly. On the other hand, the public school and community college agencies create files containing appropriate enrollments and provides access to FETPIP staff. The linkage with these files is conducted by FETPIP staff.

Expenses will have to be accounted for. Whether follow-up by record linkage is to occur through existing or additional staff will have to be determined. Costs associated with computer linkages, editing, and report generation will have to be covered. In Florida, FETPIP costs are covered through state funds made available to the Department of Education. The Texas operation is funded through contributions by agencies who are partners in the follow-up process. Another option may include building wage record follow-up into the evaluation scheme that supports a state's school-to-work program with earmarked funding for the activity.

Developing Interagency Agreements

In all probability, there will have to a formal written agreement between each agency involved in the exchange of information for computerized follow-up. With regard to the transfer of student records, the agreement may take the form of a "Buckley Agreement".[11] A Buckley Agreement will stipulate that the recipient of the individual student records (whether obtained on paper or electronic media) understands the requirements of the Amendments regarding the confidential nature of student records including conditions, prohibitions, and penalties associated with public release of individually identifiable information. The unemployment insurance agency will require a more extensive agreement.

In developing any binding agreement between government agencies, legal consultation will be required with the attorneys of the affected agencies. This consultation

will be facilitated within the "match maker" agency if legal staffs were involved early in the process. Each agency may have "boiler plate" agreements that serve as templates for contracts. There may also be template agreements that facilitate the exchange of personally-identifiable information. There will be state-specific and agency-specific assurances and certifications that have to be included. Templates and unique requirements notwithstanding, there are some basic elements of formal interagency data matching agreements that have been successfully negotiated.

The Computer Matching and Privacy Protection Act of 1988 was intended to guide federal agencies in conducting computer matching programs. It provides guidelines to agencies about the types of agreements they must have prior to conducting an interagency match. The law is only applicable to federal agencies, it also excludes the type of record linkage discussed herein from the definition of a "computer match". However, the guidelines in the law represent a good starting place in designing an agreement.

FETPIP has incorporated the following elements into its record linkage agreements:

INTRODUCTION. An introductory section should identify the agencies involved in the agreement. It should establish that each agency has certain data resources that when combined can help to achieve a specific purpose, in this case to provide information that helps to evaluate the post program employment outcomes of vocational education programs.

PURPOSE AND JUSTIFICATION. If there is a statutory basis for the computer matching program, it should be cited here. This may be a state vocational education accountability law. If the effort relates to the accountability requirements of the Carl D. Perkins Vocational and Applied Technology Education Amendments of 1990, it could be cited.

Regardless of the statutory basis, this section should elaborate on the overall purpose stated in the introductory section. It should indicate how the purpose will be met. For example, "...by generating aggregated statistics that will be used in describing the employment experiences of students after participating in a training program".

It may be desirable to include statements that relate to the cost-effectiveness of collecting follow-up data by computer matching rather than by conventional means.

IMPLEMENTATION AND RECORD EXCHANGE. These sections should include information that outlines how the computer match will be initiated. For example, the unemployment insurance agency will be contacted during a certain period regarding the volume of student records and expected time frames of the submission. A statement should outline how the data are to be exchanged, i.e., by providing computer tapes, cartridges, or accessing electronic files.

A statement should indicate exactly what data will be provided by the agency

desiring the match (for example, a file containing unduplicated Social Security Numbers) and what data will be provided back for each response record by the unemployment insurance agency.

SECURITY, PRIVACY RIGHTS, AND PUBLIC RELEASE. There should be a section or sections that acknowledge applicable state and federal laws concerning the privacy of student records as well as restrictions on the use of unemployment insurance records. This may include acknowledging that the records of the involved agencies are being used for purposes other than that for which they were designed. It should outline how the unemployment insurance agency (or the agency that is matching against student records) will handle files containing student records. It may include language that requires that they do not duplicate either the student record file or the file that results from the match for purposes other than accomplishing the match. It may require that those doing the match not examine records in the file unless it is necessary to resolve computer processing problems. The sections should require that when the job is complete, duplicate tapes or files be disposed of or purged. The sections should stipulate that while the subject files are in the possession of the matching agency, they be retained under secure conditions. There may need for some elaboration on what the secure conditions are.

If there are requirements that stipulate thresholds of aggregation for the release of data resulting from individually identifiable data, they may need to be included in these sections.

ADDITIONAL INFORMATION. Other sections will define the duration of the agreement, how it may be renewed, conditions of termination, reimbursement of costs, and authorizing signatures.

Carrying Out the Interagency Agreements

The "match maker" will be under an obligation through the Buckley Agreement to disclose only that information which is necessary and proper to accomplish particular tasks. In matching with the unemployment insurance system, only the Social Security Number is to be used. Therefore, the tape or electronic file that is to be used for the match should contain only Social Security Numbers. Files that contain detailed demographic or program data will be retained (and kept in a secure environment) while the matching process is occurring.

Chances are good that the agreement with the unemployment insurance agency will contain specific language regarding how costs for the activity will be calculated and recovered. This provision will be based on the volume going in (i.e., the number of Social Security Numbers going in) and the volume going out (i.e., the number of response records resulting from SSN "hits"). To limit unnecessary costs (and eliminate potentially erroneous "hits"), it is wise to first eliminate Social Security Numbers that are not useable.

Unuseable numbers include 123-45-6789, 111-11-11111, etc. Simple computer programming language can be developed to look for, and eliminate these from the records. Additionally, the U.S. Social Security Administration publishes monthly lists of numbers that have been assigned. The lists include a description of assignment parameters that can also be included in a computerized edit. A visual edit can help to uncover groups of numbers occurring in suspicious sequences that one may wish to eliminate from the process as well.

It will also be wise to "unduplicate" Social Security Numbers from the file being sent to the unemployment insurance agency. Duplicates most often occur because students participate in more than one vocational program. Duplicates can also occur when students register using someone else's Social Security Number. In "unduplicating" the numbers it may be appropriate to examine the names of students associated with the duplicate numbers. If the volume of Social Security Numbers is large, this can be facilitated electronically by matching on names or parts of names. If names are inexplicably different, they should be eliminated as any information gathered by the matching process would be suspect. Explainable differences include females with the same first name but different last names (because of marriage or divorce), obvious typographical errors, and the use of initials or nicknames.

When the response records are returned from the unemployment insurance agency, they will have to be eventually combined with the original files from which the Social Security Numbers were removed, edited, and unduplicated. It is important to keep in mind in the analysis process, that there will always be a larger number of jobs than numbers of people employed. In developing reports, jobs information will have to be clearly distinguished from people information. Further, there will be more program participants than people because of multiple program enrollments. Employment data on particular programs should not be aggregated across programs without recognizing this.

Security arrangements that are spelled out in the matching agreement should be followed zealously. Tapes, disks, files, and reports that contain individually identifiable information should be retained only so long as they are needed. If files are being maintained over relatively long periods because of longitudinal interests, there should be periodic justification of their retention among participants in the agreement. Records with individually identifiable data that are being retained for any use should be kept under lock and key with limited accessibility when they are not in use. Files being stored in mainframe computer facilities or personal computer hard drives should be subject to stringent security and be accessible only to authorized personnel using secure passwords.

It is important to recall that the interagency agreement allowed the "match maker" to "borrow" certain records that were "loaned" by individuals and employers to another agency. While the formal agreement will stipulate many of the circumstances regarding data handling and processing, there should be a sense of trust developed that goes beyond it. This requires working with the agencies to assist them in recognizing and resolving problems that are uncovered in the matching process. It also means continually

working with them to ease processing burdens associated with the matching effort. Where there are questions regarding the release of certain data or analysis, even if the release appears to fall within the agreement's parameters, the cooperating agency should be consulted.

One such question may arise when there are circumstances that result in requests from users regarding access to the files that resulted from the match with the unemployment insurance agency. Generally, the transmittal of files or reports containing individually identifiable records to organizations or individuals other than those that are party to the agreements should be discouraged. If it is necessary to transmit data that retains its individual character for purposes that promote the use or improvement of the data, those who are party to the agreement should be consulted. If such a transfer is agreed to, it is appropriate to remove names, Social Security Numbers, and any other information that could result in personal or specific employer identification from the records being transferred. The personal information could be replaced by dummy codes or names. In any case, an agreement similar to that described previously should be negotiated.

The original reason for pursuing a matching agreement may have been to develop state level analyses or information related to performance standards. However, once the matching data are received and the "match maker" is comfortable in working with the data, a major service can be provided to local education agencies whose information systems provided the original student data used in matching. This "service" would take the form of tailoring a variety of reports and analyses for local use. In initial efforts, this is particularly important, for the local agency may become a fan which can assist the "match maker" in working with additional customers.

Pursuing a Vision

Those who try this approach for the first time will have to decide on subsequent development and direction for the follow-up data collection effort. If the program began by involving a few localities or vocational education agencies, the next stage of development may be directed to expanding the program to include other, and eventually all, vocational education agencies statewide. Several other areas are suggested in the following discussion that may help in cultivating ideas for the future direction of a statewide computer matching follow - up effort.

1. Expand Coverage. The unemployment insurance data base does not cover all of the employment that may be available to former students. Nothing, other than the characteristics information accompanying the student's record, is known about the people who lacked a response record from the unemployment insurance wage file. A set of objectives could be established and pursued to find out more about those "not found".

While the state unemployment insurance systems are quite comprehensive, not all employment opportunities are covered. Among those not covered are federal, postal service, military, railroad, and certain small establishment including self-employment.

Additionally, employment, where commissions form the basis for remuneration are not included. Earlier in the article, employment in the military, federal career, and postal services were discussed as reasonably accessible sources of information. Collection of employment data from the unemployment insurance agencies in neighboring states is an additional option.

Not all former students will be working. Some will be pursuing further educational opportunities. It may be possible to work with the public postsecondary agencies and institutions in other states to develop matching agreements that would assist in identifying those students who are continuing their educations outside of their home state. Unemployed students might be identifiable through unemployment insurance claims records. Former students who are unemployed but are receiving some form of public assistance may be identifiable through the state public assistance agency or the agencies responsible for JTPA and/or JOBS. Some students may even be incarcerated in state correctional facilities.

2. Improve Data Quality. As has been suggested earlier, the "match maker" is somewhat at the mercy of the agencies supplying student and wage records. If elements in student records that describe socio-economic characteristics, demographic characteristics, or educational programs are not consistently recorded or reported in most cases, they will be useless in reporting results. The "match maker", however, will have a unique view of student and wage data that is not available to the agencies who provided it in the first place. Through the Social Security Number edit process, a service can be provided back to agencies that could help them improve their collection and reporting. By recognizing gaps in the records, missing items can be identified that may be desirable for program evaluation. As more schools and institutions are brought into the process, nuances in the reporting of program codes can be identified and rectified.

3. Develop Common Indicators. This article began by recognizing that those involved in public vocational education and job training programs are increasingly being held accountable through performance standards and other valuative measures. The standards and measures focus on employment outcomes to a large degree. There are additional areas where employment outcomes could be a feature of accountability efforts as well. These include high school academic and general education and university level education programs. Both public and private programs should be considered. In other words, it may be appropriate for the "match maker" to consider and pursue customers other than those associated with public vocational education or job training.

By matching groups of former students, program clients, or program participants against the same administrative data resources, a set of common indicators could be developed. The indicators may include rates of employment, rates of earnings, and employment distributions across industries. They could be broken down by age, race, sex, and certain socioeconomic characteristics. The system would accommodate state as well as local level reporting. This type of data would facilitate evaluating the employment impacts of various public and private programs and combinations of programs on some common grounds. The data could be used as the basis of a

consumer information program using information that was gathered in a common manner, driven by common definitions.

Conclusion

The purpose of this article was to provide some guidance to states who may be considering an interaction with their unemployment insurance wage record system to better understand employment outcomes resulting from employment, education, and training programs. The article is intended to be promotional in that it is written from the point of view that follow-up data collection done by linking student/participant records to wage records is a good idea. It is cost-effective in that in uses existing government sources of information that, while originally designed for a different purpose, can provide valuable insight into employment outcomes resulting from employment and training programs. It can provide good, comprehensive, understandable data at a fraction of the cost associated with traditional follow-up methods.

The article is also intended to instill a keen sense of caution in those considering the approach. Record linkages of the sort that were described must be conducted only when adequate provisions are made to protect the privacy of the individuals whose data are contained in the various systems. This means that the operational environment through which the data are collected, stored, and manipulated must be secure. This means that data must only be publicly available in a manner that precludes the identification of individual persons. Finally, it means that the data bases that are compiled as a result of record linkages should only be used in conjunction with research and evaluation activities using aggregate data.

Endnotes

1. A series of studies commissioned by Florida's Governor in 1982 were published under the general title of The Vocational Education Study. One monograph in the series, "Analysis of the Labor Market Survey" dealt with the conduct of student and employer follow-up by Florida school districts and community colleges. It indicated that both districts and colleges had considerable difficulty in obtaining satisfactory survey response rates. The Vocational Education Study recommended that a pilot effort be begun immediately to examine the feasibility of replacing traditional surveys with use of the unemployment insurance wage report data base.

2. In November 1991, FETPIP was awarded a "Davis Productivity Award" by Florida TaxWatch, Inc. In the citation accompanying the award, TaxWatch credited the program with tax savings in excess of $3.1 million per year to school districts and colleges.

3. Match rates vary according to the application. For example, match rates run from a high of about 88% with community college associate of arts degree recipients to a low of about 30% for recent releasees from state prisons. Other rates include 43% for high school dropouts, 71% for high school graduates, 69% for recent General Equivalency Diploma recipients, 40% for adult basic education students, and 66% for JTPA Title IIA adult participants.

4. The JTPA survey in Florida seeks data that are not a part of that collected through FETPIP. Similarly, FETPIP collects several items that are not in the JTPA survey.

5. An initial consortium meeting was held in Kansas City on December 21, 1993 with representation from Texas, Florida, North Carolina, and Oregon. David Stevens of the University of Baltimore's Merrick School of Business attended on behalf of states that he works with as well as his own research interests.

6. Section 462 (g) of the Job Training Partnership Act required that the Bureau of Labor Statistics determine the procedures necessary to establish a national longitudinal wage record data base from state unemployment insurance systems. Section 408 of the Carl D. Perkins Vocational and Applied Technology Education Act required that the Office of Technology Assessment of the U.S. Congress review the work of the National Occupational Information Coordinating Committee in examining uses of the wage report. These examinations were assigned to NOICC in both acts.

7. A listing of Federal level administrative record linkages is available from the U.S. Bureau of the Census in its Project Link-Link.

8. Public Law 93-579, the Computer Matching and Privacy Protection Act of 1988 (Public Law 100-503) helps define some of these purposes.

9. If the connection between education and training programs and the labor market can be improved through the use of follow-up data, there is an argument that can be used to connect that improvement to reductions in unemployment. Reductions in unemployment and/or the length of unemployment result in reduced employer liability necessary to support the system.

10. State practitioners will have to determine what is feasible in this pursuit. It may be appropriate to begin at a relatively modest level by working with a particular institution. On the other hand, if statewide data are available, it may be appropriate to begin at the state level.

11. The Buckley Amendments amended certain provisions of the Privacy Act to assure that special protection and security be afforded to the confidentiality of individual student records.

Bibliography

Anderberg, Marc and Pfeiffer, Jay "Proposed Interstate Consortium Regarding the Use of Record Linkage Technologies in Education and Training Follow-up" (unpublished-circulated among targeted states for reaction), Austin. December 1992.

Florida Education and Training Placement Information Program, Annual Report: August 1990, Florida Department of Education, Tallahassee. August 1990.

Harris, Linda et. al. Vocational Education Study: Analysis of the Labor market Information Survey, Office of the Governor, Tallahassee. July 1982

National Commission for Employment Policy, Using Unemployment Insurance Wage-Record data for JTPA Performance Management, Research Report Number 91-07, Washington, D.C., June 1992.

Partners in Productivity, The 1991 Davis Productivity Awards: Winners in the Spotlight, Florida Taxwatch, Inc. Tallahassee. November 1991.

U.S. Department of Commerce, Bureau of the Census, "Federal Level Linkage Activity Report" in written communication to FETPIP, Washington, DC, March 1994